DOCUMENT RESUME

ED 221 557 TM 820 578

AUTHOR Burstein, Leigh
TITLE Investigating Social Programs When Individuals Belong to a Variety of Groups Over Time.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO CSE-R-173
PUB DATE 81
NOTE 48p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Early Childhood Education; Educational Change; *Group Membership; *Instructional Innovation; Models; *Program Evaluation; Research Design; *Research Methodology; Research Problems; School Organization; Social Structure; *Student Development
IDENTIFIERS Planned Variation; Project Follow Through; Project Head Start

ABSTRACT

Research on school reform efforts and reform evaluation and the recognition that individual students are in constant transition and are members of multiple groups whose dynamics differ over the duration of schooling were examined. A conceptual framework for developing design and analysis strategies for investigating educational reform efforts is described. A research strategy of examining educational processes and effects generated by reform programs on outcomes is discussed. A setting in which several well-defined programs of early education are implemented in multiple sites for a period of several years is examined. Suggested investigation criteria in research and evaluation in school reform include: (1) identification of interactive effects of program types with types of students; (2) the relation of educational processes to performance and program definition to processes; (3) the degree of program implementation precepts in relation to estimates of program impact; (4) group composition and student role effects on individual opportunities and behavior; and (5) students' class to class transitions and program/non-program participation needs to determine discontinuities in instructional and social experiences for educational performance. (CM)

INVESTIGATING SOCIAL PROGRAMS WHEN INDIVIDUALS BELONG TO A VARIETY OF GROUPS OVER TIME

Leigh Burstein

CSE Report No. 173
1981

Center for the Study of Evaluation
Graduate School of Education, UCLA
Los Angeles, California 90024

The project presented or reported herein was performed pursuant to a grant from the National Institute of Education Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

The basic charge for this paper indicates two operational purposes. First, perspectives on the documentation of school improvement efforts are to be addressed. Philosophical, methodological, and practical considerations are supposed to guide this examination. The extent to which these concerns can be satisfied simultaneously remains to be seen. Any specific methodological approach reflects, at least implicitly certain philosophical perspectives (e.g., about the nature and investigation of cause; about valued educational processes and outcomes) as well as notions about the practical limits on empirical inquiry (e.g., reasonableness of intrusion in normal educational routines, invasion of privacy, ability to measure behaviors of interest accurately, feasible sample size and study duration). While it is difficult to make them explicit, it is hoped, nonetheless, that the philosophical and practical considerations that both govern and constrain the methodological considerations will be evident.

The Transitory Character of Educational Experiences

The second purpose is to consider the implications for research on and evaluation of school reform efforts of the recognition that during their schooling experiences, individual students are in constant transition and are members of multiple groups which differ over time. It takes no special wisdom to recognize that students develop and learn in a variety of social and educational settings during their school years. The potential demands on children to adjust to new roles as they shift educational and social groups are as much a recurring event in the growth process as the need to develop

their cognitive abilities, attitudes, and self-identities. The dynamic nature of child development and the "natural transitions" involved in changing group membership, then, are normal features of children's lives. As such research efforts to document the developmental process, including the role of "natural transitions", are the normal fare of social scientists interested in those aspects of development and socialization (e.g., learning, attitude formation, peer-group relations, status identification) that occur in educational settings.

What would seem to make the above issues (i.e., students being in constant transition and being members of a variety of groups) of interest in research and evaluation of school reform efforts is their implications for the design and study of these efforts. Specifically,

1. School reform entails change, a departure from ongoing practice. Profound, and thus potentially important, departures necessitate substantial transitions for students and other participants (teachers, principals, parents, etc.) in the educational process. The effects on children of the "natural transition" process may pale by comparison. The impact on children's lives returns to a state of equilibrium only after the reform activities become the educational norm.

2. Specific reform efforts may blend smoothly with the educational and social systems in schools or may work at cross-purposes. Those reforms that least disturb the students' roles within the groups to which they belong are least likely to disrupt developmental and socialization patterns.

3. Regardless of the focus of the reform and its sensitivity to the existing educational and social system, its impact cannot be determined by research and evaluation methods which ignore the dynamic properties of the change process. These dynamic properties are fostered by the contact of the reform effort with the "natural transition" of pupils through schools and with the pupils' roles within a variety of groups over time.

Working Definition of School Reform

The above is stated rather abstractly. The characteristics of the school reform under discussion have not yet been identified. Nor have the functioning meanings of "natural transition" processes or "membership in multiple groups over time" been provided. At this point, we attempt to bring the above concepts into better focus by choosing first a specific kind of school reform and then describing how pupils' "natural transitions" and group memberships operate within schools implementing such reforms. The eventual intent is to provide a conceptual framework for the educational experiences of children in such settings and thereby provide a basis for developing design and analysis strategies for their investigation.

Multiple Well-Defined Programs

The kinds of school reform efforts to be considered are educational interventions like those in operation under Planned Variation Head Start and Follow Through Programs (Rivlin & Timpane, 1975). These interventions involve the implementation within school systems of innovative programs of pre-school and primary education. The programs (approaches/models) are allowed to differ in

educational philosophy, curricular emphasis, and the importance they place on various cognitive, social, and psychological goals.

We further assume that the various approaches represent well-developed educational offerings that are grounded in theories of development, instruction, and learning. Moreover, it is assumed that the approaches can be taught, with perhaps differential success, to teachers who will attempt to implement them in pre-school and primary classrooms. While special efforts might be made to ensure the fidelity of the implementation at the initial stages of the reform, eventually the programs would be expected to be developed to the degree that they could be disseminated intact to any number of classrooms and schools.

A few caveats about the assumptions in the above paragraphs seem warranted. We are not so naive that we believe that past and present efforts at school reform ever operate as neatly as assumed above. For a variety of reasons, some of which are covered in this paper, there is a great deal of slippage from intent to development to implementation to dissemination. Nonetheless, we would be hard-pressed to think of any significant educational innovation attempted in the last 25 years -- the whole spectrum from post-Sputnik curriculum projects through the various Great Society programs (Head Start, Follow Through, Title I, etc.) to bilingual education, PL94-142 special education reforms and school desegregation -- that was founded on any less of an ideal view about the possible consequences of programs of educational and social change.

There are simply enough problems inherent in examining the effects of well-defined approaches confronting an existing educational system and social network without taking on the additional problem of what to

do when the intervention itself is ill-defined. Evaluations of the effects of Title I, bilingual education, PL94-142, and school desegregation cannot avoid considering ill-defined interventions. The conceptual framework offered below for examining well-defined interventions may still apply. However, interventions mirror existing educational practice to the degree that the constraints on the intervention (i.e., explicit goals, and program definition and practices) are removed. As such, ill-defined interventions may best be treated as simply variants of standard practice and investigated accordingly.

There is a more pragmatic justification for the constraints imposed on the kinds of school reforms under consideration. Our view of the impact of reform efforts is that they can potentially change children's educational experiences. Unless one starts with well-defined intervention, it is difficult to distinguish within-program heterogeneity from normal variation in educational practices. Well-defined interventions at least have the possibility of being distinct from typical practices. Only under such circumstances can one begin to expect new programs to have identifiable consequences.

Multiple Sites

Other aspects of our working definition of school reform include the availability of multiple sites for each program variation wherein students are program participants. With data from multiple sites per program, variation in program implementation can be investigated. It is also possible to study interaction between program approach and the setting in which it is implemented under these conditions. There is certainly sufficient indication from past evaluations of Follow Through

(e.g., Anderson et al., 1978; Haney, 1977; Kennedy, 1978; Stebbins et al., 1977) that variation in program implementation and interactions with educational settings are integral features of school reforms and thus warrant investigation as part of future research and evaluation activities.

Multiple Years

Well-defined school reform programs are typically implemented in multiple grades. Moreover, it usually takes several years for the program to become properly implemented. It is also the case that students seldom remain in a reform program for their entire schooling experience.

Given the above, it is reasonable to presume that a multiple-year study of the effects of program participation is warranted. Ideally, it would be best to collect information about student experiences during several years in the program and for some years after leaving the program. Recent reports on the lasting effects of early childhood educational interventions (Brown, 1977; Lazar et al., 1977; Palmer, 1977) show the potential value of continuing investigations of program effects long after students leave the programs. We see no reason to ignore such possibilities in planning future research and evaluation of school reform efforts.

If future programs mirror present ones in terms of diversity of approaches and sites and program duration, then several proposals for multi-year research, development, and evaluation studies of these reforms already exist (e.g., Ellett, Haven, Pool, & Smock, 1979; Weikart & Banet, 1975). The plan proposed by Ellett et al. (1979) is sensitive to both multi-site and multi-year features of programs.

Under their plan, each approach would be implemented in a number of sites (communities, schools, etc.) and the study would follow several cohorts of students from kindergarten (K) through grade 6. Since the approaches are implemented only in grades K through 3, the study follows students beyond the end of their direct contacts with the program. Later, we shall discuss a shortened version of the Ellett et al. plan only for the sake of detailed illustration of study features within a shorter time frame.

Multiple Outcomes

Because specific programs start from different perspectives, the kinds of specific outcomes of interest of each program tend to differ, often drastically. Nonetheless, all are trying to improve the educational and life chances of children (usually poor children). In the broadest sense of the term, there are common types of child outcomes of interest. These outcomes include educational achievement (both short-term and long-term), attitudes towards self and schooling, initiative, independence, adaptability, school attendance, special education placements, grade retentions, and "psychological well-being" (Scan Sponsor Task Force, 1980; Haney, 1977). The full set of outcomes is often referred to as the development of educational and social competence. At a variety of points during and after membership in classrooms supposedly implementing one of these programs, program effects on some subset of these outcomes are to be assessed.

There has been substantial discussion (see e.g., Ellett et al., 1979; Madaus, Airasian, & Kellaghan, 1980; Rivlin & Timpane, 1975; Wargo & Green, 1978; Weikart & Banet, 1975) about whether programs

with diverse objectives should be examined only with regard to outcomes derived from their own objectives or whether a common set of outcome measures should be applied to all programs. While it is not the purpose of this paper to argue for any general strategy for selecting outcome measures, several comments seem warranted. First, we view the purpose of school reform research and evaluation to be the generation of evidence about the likelihood of achieving a wide array of educational and social outcomes from the programs under study. Thus, measures representing the full range of outcomes of interest ought to be obtained from participating programs.

At the same time, we foresee the possibility that individual programs might choose to measure specific types of outcomes in different ways. This would occur simply because given instruments overlap to different degrees with a specific program's goals and instructional practices (e.g., Armbruster et al., 1977; Porter, Schmidt, Floden, & Freeman, 1978; Walker & Schaffarzick, 1974). Given this latter concern, it seems reasonable to allow programs to select instruments that are nominally comparable in measuring a desired outcome rather than requiring strictly common measures across programs.

In practice, it is perhaps more desirable to have specific programs nominate a set of measures which they believe to be capable of measuring their program's impact on the range of outcomes of interest. Then, as suggested by Rivlin and Timpane (1975, p. 13), each program could be compared with other programs (as well as with non-participants) on its own specifically chosen measures and on a set of measures chosen to represent the goals and practices of other programs. Such a

strategy would yield evidence about a given program on the basis of its own objectives and in terms of the objectives of other programs. With this type of information, potential users can apply their own values in weighting various measures to choose among program alternatives, a decision strategy consistent with our view of the purpose of the proposed studies.

Multilevel Outcomes

In the final analysis, most school reform efforts are directed toward changes in the educational and life chances of the child. Thus, it is reasonable to concentrate study efforts on the educational experiences and outcomes of children in the presence of school reforms.

However, overall concerns about pupil processes and outcomes do not mean that concerns about instructional and schooling practices should be excluded (Burstein, 1980b; Haney, 1980). On the contrary, investigations of the behaviors of classes, teachers, and schools are essential if the character and consequences of school reforms are to be documented. Behaviors of higher-level units (teachers, classes, schools) are both intermediate outcomes in studies of program implementation and effects and antecedents to pupil processes and outcomes.

Moreover, there is much to be learned about the inherent tradeoffs within and among specific programs from examining the distribution of pupil processes and outcomes within the larger units (e.g., Bidwell & Kasarda, 1980a, b; Bossert, 1979; Brown & Saks, 1975, 1980a, b; Burstein 1980a, b). For instance, it is quite conceivable that programs that emphasize cooperative learning arrangements result in different arrays of pupil experiences and outcomes than more competitive arrangements (e.g., Sharan, Ackerman, & Hertz-Lazarowitz, 1979-80). Focussing on the distributions of

experiences and outcomes for groups (classes, schools) may provide clearer evidence of the consequences of alternative resource allocation decisions than strictly individual-level analyses. And, since no penalty accrues from conducting analyses at multiple levels, there seems to be no good reason to restrict the investigation of school reforms to any specific unit or level.

The Role of Educational Processes and Experiences

So far we characterized the study setting as one in which several well-defined programs of early education are implemented in multiple sites for a period of several years. Furthermore, though the various programs are allowed to vary in philosophy and interest, each presumably fosters educational changes which affect experiences and outcomes at the multiple levels (pupil, teacher, class, school, etc.) of the educational system. In essence, the nature of the reform is to introduce programs offering a specific array of activities and possible experiences into an already dynamic educational and social system to modify the system's equilibrium toward more positive experiences and outcomes. When the reform is targeted toward specific segments of the study population (e.g., poor children), it is hoped that the program will help to ameliorate differences in educational and social competency that distinguish the target students from the general student population.

Under the conditions described above, it should be evident that regardless of intent, school reforms like Follow Through should not be viewed as static interventions into normal educational routines.

Programs cannot be characterized simply by their labels or definitions. They represent an intended array of educational processes confronting existing educational and social arrangements.

To arrive at any reasonable methodology for investigating programs like Follow Through, it is necessary to redirect interest away from program labels and toward program processes, both intended and actual. The investigation then evolves into a two-stage examination of the effects of school reforms on students, teachers, schools, and communities. In one stage the educational and social experiences (which we call "educational processes") of students (teachers, etc.) participating in specific programs are examined. For example, if program A emphasizes small-group instruction, then a series of questions about such matters as (a) whether pupils participating in the program receive more small-group instruction than non-program students, (b) do students stay on-task (experience more and/or higher-quality teacher-student and student-student contacts, cover more content, receive more appropriate instructional materials and so on), and (c) what additional resources and other modifications are required to make the shift to small-group instruction would be studied.

In the other stage, the effects of the educational processes generated by the reform effort on the outcomes for students, teachers, classes, schools, and communities are investigated. A possible question would be whether the educational processes engendered by the small-group instruction under program A affected the educational and social competencies of children. For instance, one might look for small-group instruction to lead to greater student-student contact which in turn results in more cooperation among students,

enhanced self-esteem and improved ability to handle intellectually more complex cognitive tasks (e.g., Calfee & Brown, 1979; Johnson, 1981; Sharan, Ackerman, & Hertz-Lazarowitz, 1979-80). The intent here is to examine the direct linkages of the educational processes to the outcomes.

The two-stage research strategy does not necessarily include direct effects of specific programs on educational outcomes. Program effects on outcomes may be strictly indirect through the educational processes they generate. In one sense, programs serve as moderator variables by making the occurence of a given array of educational processes more or less likely. "Direct" effects of specific programs on outcomes might be believed to occur when (a) certain programs more consistently foster a given educational process (such as small-group instruction) than other programs and (b) the specification of educational processes in the second-stage analysis is incomplete (e.g., when the instructional and social processes associated with small-group instruction are not measured adequately). However, even these presumed direct effects are spurious since the programs are simply serving as proxies (indicators) for the inadequately measured educational processes they generate.

Our stance in favor of a two-stage research strategy is in line with a variety of investigators who call for general inquiries into the effects of social programs (e.g., Cohen, 1975; Cronbach & Associates, 1980; Weiss, 1977) or for studies of program processes and implementation (e.g., Lukas, 1975; Stallings, 1975; Weikart & Banet, 1975). Moreover, rather than picking sides in the current debates on whether to focus strictly on program processes or on program outcomes, we choose to

consider both but only the understanding that educational processes are key elements of the investigation. Educational processes are the "medium of exchange" in the school reform effort. A particular program implemented in a particular educational and social setting generates educational processes which are then used to "purchase" educational opportunities and outcomes for program participants. The program is then advantageous to the degree to which it can generate processes of high value in purchasing desired outcomes.

Patterns of Effects from Well-Defined School Reforms

Later on, we will provide more detailed accounts of possible design and analysis approaches for the two-stage research strategy discussed above. At this point, we present several possible explanations of the consequences of existing preschool and primary interventions such as Head Start and Follow Through. These explanations result from attempts to understand the findings from various early education evaluations (e.g., Anderson et al., 1978; Baker, 1976; Cicirelli et al., 1969; Cline et al., 1974; Haney, 1977; House et al., 1978; Lazar et al., 1977; Palmer, 1977; Smith, 1975; Stebbins et al., 1977).

None of the specific explanations are new. They can be found in various sources including Cronbach and Snow (1977), Cronbach and Associates (1978), the full set of papers in Rivlin and Timpane (1975), and the report of the Scan Sponsor Task Force on Follow Through (1980). We are attempting to combine various explanations in order to understand why the literature on these programs has been so confusing and what it might take to achieve greater coherence of findings from the next

round of Follow Through research and development. If the latter is to be accomplished, then explanations must point to a general conceptual framework for investigating the consequences of well-defined school reform efforts. We shall offer a possible framework to guide such investigations in the next section.

It is necessary to state the findings about the effects of early education interventions before one can begin to understand explanations of how they occur. There are a number of both primary and secondary sources from evaluations of Head Start and Follow Through. Without recounting the complete list of findings or sources, one can still get the flavor of the results from the ones given below. With respect to Head Start and other preschool programs, the literature indicates:

(1) For the original set of Head Start programs (prior to 1969), preschool interventions had immediate favorable impact on both cognitive (e.g., IQ) and other measures (see e.g., Cicirelli et al., 1969; Datta, 1975; Rivlin & Timpane, 1975).

(2) The early gains faded without additional intervention upon entry into regular school and scores declined after the third grade (same as for (1)).

(3) In the Head Start Planned Variation study, in which Head Start programs adopted curriculum models used in the Follow Through Program, children's test scores increased substantially on all outcome measures (Smith, 1973, 1975) at the end of the Head Start experience.

(4) There were strong differences among Planned Variation models in effectiveness though no model stood out as more or less effective than the others on most of the outcomes (Smith, 1973, 1975).

(5) Certain models had strong positive effects on certain outcomes. For example, approaches that encourage structured academic emphasis and drill on cognitive tests were particularly effective at imparting information that is easily taught through systematic drill. The High/Scope Model was more effective in raising Stanford-Binet scores than other approaches (Smith, 1973, 1975).

(6) Recent studies of the lasting effects of a number of "carefully conceived and implemented" early education programs indicate that program participants were less likely to be retained in grade or to be assigned to special education classes, and had higher arithmetic and reading achievement, and IQ scores. These results were found for students in grades 3 through 8 and from a variety of approaches. Some programs show progressive relative improvements on cognitive tests over time while others had constant or no gains relative to comparison children (Lazar et al., 1977; Palmer, 1977).

With respect to Project Follow Through, the studies indicate:

(1) The short-term effects of Follow Through were positive and small for all outcomes (Cline et al., 1974; p. VII-11ff).

(2) Sponsor (i.e., specific programs/approaches/models) diversity was great (Cline et al., 1974, p. VII-11ff).

(3) According to later reports, the effectiveness of each Follow Through model varied substantially from site to site; differences in between model averages were small in comparison (Anderson et al., 1978; Haney, 1977; House et al., 1978; Stebbins et al., 1977).

(4) Models that emphasize basic skills succeeded better than other

(5) Students participating in models that did not place their primary emphasis on basic skills, fare less well in basic skills tests than non-Follow Through Children (same references as in (3)).

(6) Some models are more successful in their most disadvantaged sites (same references as in (3)).

(7) Most models are more effective during kindergarten and first grade than during second and third grade though the effects for some models grew over time (see Haney, 1977; Stebbins et al., 1977).

Explanations of the findings from past early education studies are derivable from various social science perspectives on the behavior of individuals over time in naturally varying social settings and in the presence of innovations introduced into the social settings. In the present case, the major explanatory mechanisms would seem to be the following:

- When properly implemented, well-conceived programs with developed curricula and training procedures affect those behaviors suggested by the underlying theories upon which the programs are based.

The above explanation simple credits well-defined, theoretically conceived programs with the ability to accomplish the outcomes they are designed to deliver. For example, it is clearly the case that the basic skills orientation and curricular emphasis of the Oregon Engelman-Becker model and the Kansas Behavioral Analysis Approach were successful at improving basic skills performance in the Head Start and Follow Through evaluations. Likewise, the success of children participating in the High/Scope model programs on measures of IQ (Lazar et al., 1977; Smith, 1975) and measures of achievement in later grades (Lazar et al., 1977;

Palmer, 1977) contrasted with poor performance on measures during the years of Follow Through participation (Stebbins et al., 1977) might be viewed as evidence that the High/Scope program fosters its intended goal of complex general intellectual development and skills (better reflected in later years results) rather than basic skills mechanics (reflected in the Follow Through Test Battery).

According to this explanation, the results for the Follow Through models labelled as "Affective-Cognitive" by Stebbins et al. (1977) can be interpreted in several ways. Perhaps these programs were, in general, improperly implemented so that past studies provide inadequate tests of their presumed benefits (Stallings' (1975) evidence is equivocal on this point). Another interpretation might be that expected consequences of these programs for children's behavior were not clearly delineated nor understood. An explanation proffered by program proponents is that the actual measures used in earlier evaluations were inadequate to measure the expected outcomes of these programs or that it was too early to tell.

Our support of the notion that programs can accomplish their intended goals does not rely solely on Head Start and Follow Through evidence. There are too many indications from other curriculum projects to rule out the plausibility of this assertion. It appears that evidence counter to this explanation is generally traceable to one of several phenomena. First, the possible consequences of a particular innovation tend to receive less careful attention than the creation and design of the innovation itself. And, even when consideration has been given to consequences, the range of outcomes considered is typically much narrower than those likely to occur when an innovation is introduced in an on-going social system.

Second, little attention is given to questions of the range of expected effects of an innovation. The innovator seldom asks for whom and under what circumstances the innovation might be expected to work. Consequently, estimates of program impact typically gathered may, by happenstance, be based on cases (persons, sites) where the innovation is not likely to demonstrate its benefits.

Finally, most studies have been deficient in matching the instrumentation (both its content and timing) to the attributes and expected consequences of innovations. Despite the apparent sophistication in both state-of-the-art instrumentation work and innovation design, we are still novices at achieving adequate correspondence between the two. Under such conditions, failure to find evidence of program impact can be simply the result of mismatches.

- Specific programs are more suited for some students than for others. Individual differences among students interact with program characteristics to yield differential outcomes.
- Specific programs are more suited for some teachers (settings) than for others. Individual differences among teachers interact with program characteristics to yield differential outcomes.

The second and third explanations for patterns of effects of school reforms refer to the possible differential effectiveness of innovations across individuals and settings. There is sufficient literature on the interaction between aptitudes and instructional methods (e.g., Cronbach & Snow, 1977) to warrant careful consideration of the conditions under which specific programs can be expected to accomplish their objectives.

This interaction perspective is largely absent from past evaluations of school reforms. Neither the sponsors nor the evaluators of Follow Through and Head Start models paid much attention to possible interactive effects. Though these programs were intended for poor children, the possibility that the children served differed sufficiently in ability and personality to warrant investigations of whether specific learning environments were suitable for specific children were not carefully considered. Yet, unless abilities and personalities are completely malleable, a given program could not be expected to work for all children.

The same concerns can be directed toward the lack of consideration given the matching of teacher attributes with program characteristics. Teachers surely vary in preferred teaching style, instructional skills, and preferences for specific types of children and school settings. One would not expect the same consequences from asking teachers to adopt programs either conducive or antithetical to their preferred style of operation. While it might be possible to train the teachers to implement programs different from their traditional styles, the necessary retraining should vary according to the adjustment required. Moreover, once training has been discontinued, teaching practices may tend to revert to pre-program characteristics.

The viability of these explanations for past results may be evidenced in the patterns of variation in program implementation (Stallings, 1975) and program effects (Anderson et al., 1978; Cline et al., 1974; Stebbins et al., 1977). However, since these studies do not explicitly investigate possible interactions, we are unable to rule out rival hypotheses.

There is already a shift in the early education literature toward searching for better matches between the characteristics of interventions and the characteristics of participants. In his paper on the effects of early childhood educational intervention, Palmer argues that

> the continued initiation of longitudinal studies must first answer a series of questions related to the single most important question demanding an answer with our present knowledge. What kinds of interventions are best for what kinds of children? Almost certainly there is no single program which will be best for all children regardless of region, ethnic background, and community and family environment. (1977, p. 35)

Palmer's remarks serve to emphasize the need to explore differential program effects in future research on the impact of school reform efforts.

- Well-designed and implemented early education programs socialize poverty children to the student role. As a consequence, program participants are better prepared for entering regular schools than poor children who have no pre-school experience or strictly non-educational day care experiences.

Regardless of program orientation or outcomes, participating children learn about being in a classroom setting and working with teachers and other children. This "educational" exposure prepares them, to a certain degree, for the new educational experiences they will have upon entering regular schools. Consequently, program participants adjust more quickly than non-participants to the student

role upon school entry. As a result, they are less likely to appear to exhibit "learning handicaps" requiring special treatment such as special education placement or grade retention.

If programs engender the socialization effects described above, then one would expect the proportions of program children assigned to special education and retained in grade to be lower than for non-program children. Recent reports on the long-term effects of early education interventions (e.g., Lazar et al., 1977; Palmer, 1977) find that a variety of intervention programs were successful at lowering rates of retention and special education placement. The fact that these effects held up across a variety of intervention strategies clearly points toward the likely role of school socialization as a mediating mechanism.

The positive benefits of lower retention rates and special education placements are evident from a number of perspectives. Reduced retention and special education placement may be simply the consequence of better in-class behavior and performance. Poverty children may also accrue the presumed edcational and social psychological benefits of learning in regular class settings and the avoidance of stigmatization. (The literature of pull-out programs and mainstreaming is relevant here.) Moreover, the costs of grade retention and special education, for both the schools and the student, should not be overlooked.

- Innovative instructional programs typically differ substantially from traditional instruction. Students leaving (or entering) such programs encounter substantial discontinuities in instructional

experiences. Unless programs directly affect students' adaptation skills, the discontinuity of experiences can be expected to cause program impact to decay.

The very nature of instructional innovations require that they differ from normal practice. Past Follow Through and Head Start models clearly fit this definition of innovation. For example, classrooms implementing the Oregon model emphasize strong teacher direction in whole group sessions with group response to a much greater degree than traditional classrooms. Other programs place greater emphasis than traditional classrooms on student choice and control of learning with the teacher responsible for establishing a learning environment in which students may fully exercise their choices.

The characteristic distinctiveness of such innovative programs have natural consequences. These programs may require a radical adjustment in teachers' traditional instructional styles and thus represent a discontinuity in teacher practices. More importantly, the instruction in the program is likely to be uncharacteristic of instruction in non-program classrooms, especially in higher grades where the program does not exist. Thus, students participating in innovative programs are likely to experience substantial discontinuities in instructional practices when they leave the program.

There are two ways of describing why discontinuities associated with innovative programs can adversely affect students. One way is to recognize that typical mixtures of instructional experiences within schools are not necessarily additive. That is, while teacher A in grade G and B in grade G+1 may both operate high-quality instructional

programs, students entering teacher B's class after studying with Teacher A may find the shift detrimental because of difficulties adapting to Teacher B's instructional style and classroom organization. Or, even if instruction of type B is generally better than instruction of type A, two years of type A (i.e., AA) might be better than a year of A followed by a year of B (AB) because of the discontinuity associated with the change in instruction. According to this view, discontinuities in experiences are the norm for students participating in innovative programs.

Another way of viewing the effects of the distinctiveness of innovative programs on later learning is to consider the opportunity consequences of discontinuities. The "role of student" in non-program classrooms is likely to differ substantially from the student role in the innovative program. As a result, unless the program directly fosters skills in adaptation to new settings, program participants will spend more time than non-program children becoming socialized to a changed student role. The extra time spent adapting to the new role is time unavailable for learning new material. Thus, even if program participants enter with a "knowledge" advantage, this advantage may decay due to differential time devoted to shifting to a new student role. As a consequence, non-program students have the opportunity to catch up.

The above explanation is based on the notion that, other things being equal, continuity of experiences is important for the educational and social development of children. This does not mean that experiences of children should be held constant or restricted. The concern is with major shifts in the educational and social system of children,

shifts that amount to substantial discontinuities in their everyday life. What is surprising is that educators who readily acknowledge the disruptive effects for children of parental divorce, and births and deaths in the family rarely think about the changes in experiences children undergo during early school years. Otherwise, more thought might be given to the Swedish system of keeping the same teacher throughout the first three years of schooling.

Concerns about the impact of educational discontinuities on children are not intended to serve as indictments of interventions per se. On the contrary, shifts in experiences can be beneficial for some students and obviously have been for participants in early education programs as evidenced in the long-term effects results cited by Lazar et al (1977) and Palmer (1977). We do imply that innovative programs need to be more cognizant of the importance of discontinuities and need to prepare program participants (and teachers in higher grades) to adapt to new experiences as smoothly as possible.

This set of explanations does not exhaust the possibilities in describing how school reform efforts have affected the educational and social competencies of children. Moreover, the explanations are stated in general terms. The details of certain explanations, such as how to define discontinuity of experiences or the role that the content of instruction plays throughout, have not been provided. Nor have the exceptions to general patterns been delineated or explained. Thus, both the explanations and presumed exceptions can serve as the basis for future research and evaluation of school reform efforts.

Conceptual Framework for Investigating Program Effects on Students

In generating explanations of patterns of effects, our purpose was to lay a foundation for a plausible conceptual framework to guide future investigations of the impact of school reforms. Individual explanations serve to identify both specific elements of educational and social systems in which innovations are introduced and processes that occur as a result of the innovations. The elements are the characteristics and attributes of individual students, families, groups of students, teachers, classes, groups of teachers, schools, and communities. The processes are developmental, instructional, curricular, psychological, interpersonal, and social. Both elements and processes can take on either static or dynamic properties though the latter are more likely in school settings, especially those with large numbers of poor children participating in school reform programs.

At this point, we consider directly a possible conceptual framework for investigating the effects of school reforms on student performance. This framework is intended to be sensitive to the explanations of patterns of effects from past reform efforts. It is also intended to be consistent with the two-stage research strategy wherein educational processes and experiences are viewed as both consequences of school reform programs and antecedents of educational outcomes.

A general model containing the essential elements and processes of the conceptual framework is given in Figure 1. Two years of program exposure and one year of post program schooling are depicted; more years of each could be included without loss of generality.

Figure 1

The interrelations among five distinct classes of variables are incorporated in the model: program, instruction, class composition, student entering characteristics, and student performance. Each class may represent many distinct variables (or sets of variables). For example, "instruction" refers to the various characteristics of the instruction a student receives in a specific classroom. Particular teacher attributes (e.g., warmth, enthusiasm, clarity of presentation) and instructional processes (e.g., structure, grouping, pacing, types of reinforcements, teachers' questioning behavior, quality and variety of instructional materials) both fit under the instruction rubric. Certain aspects of the instructional practices also provide evidence about the degree of program implementation. Nonetheless, any measure of program implementation would still fall within the "instruction" category for present purposes.

The term student "performance" is meant in the broad sense; the full range of educational, social, and psychological outcomes fit under this general rubric. The restriction to student outcomes could be broadened to include other units (teachers, classes, schools) but not without making the task of generating the framework even more unwieldy than it will appear here.

The role of class composition in the model is multifaceted. The overall level and heterogeneity of ability in a class places constraints on instructional content, organization, and management. The consequences of these constraints vary for different reform programs.

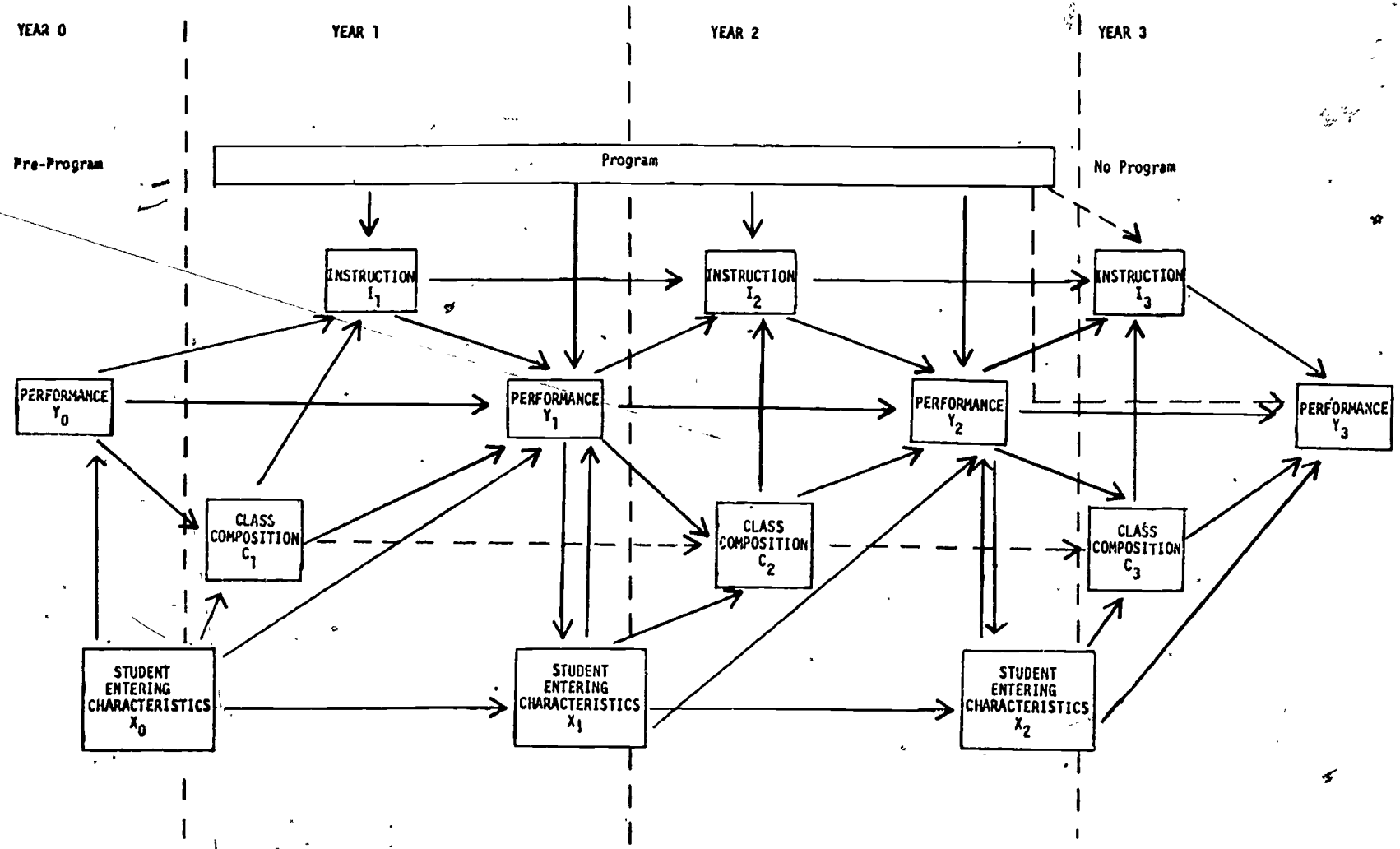

Figure 1. General model of program/instruction effects on student performance over time.

Class heterogeneity places a strain on time and resources in individually prescribed educational programs. Decisions about the pacing of instruction become more difficult in programs emphasizing large group instruction.

The student's role within the classroom is also directly influenced by composition (Burstein, 1980b; Firebaugh, 1980; Webb, 1980). There is obviously a complicated balance between having classmates compatible in ability and temperament versus having peers that are more or less able and/or have contrasting personalities. Either combination might foster intellectual, social, and psychological growth under the "right" conditions. Here, again, programs with different emphases and organization might interact differentially with class composition, making a given student's role more comfortable or stressful.

The pattern of relationships depicted in Figure 1 include the following:

(1) Student entering characteristics (ability, "preferred learning style", motivation to learn, "preparation for learning") affect performance at any point in time.

(2) Entering characteristics interact with program characteristics to give certain students relative advantages in certain programs (e.g., low ability students benefit from relatively higher levels of teacher control and direction for language and mathematics mechanics).

(3) Programs interact with teacher characteristics (preferred style, personality).

(4) Classroom composition (ability distribution, personality, presence/absence of demanding/disruptive students) affects instruction (emphasis, amount of material covered, organization).

(5) Students' shared educational and social experiences in classrooms depend on student entering characteristics, class composition, instruction, and program characteristics.

(6) Students from same class in year 1 may be assigned to different classes in year 2 or may leave the school.

(7) Students not present in year 1 may enter school (and thus program classes) during year 2.

(8) Teacher implementation of programs may differ for year 2 from year 1.

(9) Instructional characteristics (e.g., teacher "style", organization) may differ from year 1 to year 2 and effect of instruction year 1 followed by instruction year 2 is not necessarily additive.

(10) Classroom composition characteristics may differ from year 1 to year 2.

(11) Conditions (1) - (5) hold for year 2 in similar fashion as for year 1.

(12) Program differs from "normal" standard instruction and may interact.

Though instruction of Type A may be better than instruction of Type B, instruction of Type B might be better for students following participation in the program than Type A would be.

(13) Conditions (1) - (10) hold for year 3 in fashion similar to year 1 and year 2.

Implications for Design and Analysis

In theory, the conceptual framework incorporates sufficient time and features of the experiences of the students in school reform programs to investigate the explanations offered in the previous sections. In practice, the substantial number of variables that could be included within each general set, the problems of measuring each key variable, the potential for complex interrelations among variables within and between sets, and finally, the resulting analytical complexity, combine to thwart any attempt to treat the conceptual framework as a design and analysis blueprint. While the state of the art in measuring specific sets of variables and examining their relationships within and between sets has advanced from that available for the previous round of Follow Through studies, models of this size and complexity are simply not yet amenable to identification and estimation and may never be in education. Nonetheless, the framework will serve its intended purpose if it suggests questions that should be asked and the kinds of studies that might be used to investigate them.

There are at least five critical features of the conceptual framework to which the design and analyses in research and evaluation of school reforms should attend:

(1) Interactive Effects -- Investigation of reform programs should be directed toward identifying which programs work with which type of students.

(2) Programs as Educational Processes -- Rather than treating programs as simply ascribed (i.e., distinct from each other but uniformity

within program), programs should be viewed as antecedents of an array of educational processes. As a consequence, investigations of program effects should focus on (a) the relationships of educational processes to performance and (b) the relationship of program definition to educational processes. The interpretation of direct program effects on performance is problematic at best due to such problems as the sensitivity of outcome measures to program goals and possible Hawthorne-like effects. Investigations of the relationship of processes to outcomes offers the opportunity to identify the correspondence between the two, leaving determinations about preferred outcomes, and thereby processes, to decision-makers.

(3) Implementation Effects -- The fidelity of estimates of program impact on educational processes and consequently on performance, depends on the degree to which program precepts are properly implemented. Investigations of the relationships of programs to processes accomplish two purposes. First, they provide evidence on the ease with which various programs implement specific practices. Second, such investigations can suggest to decision-makers the likelihood of obtaining desired processes from specific programs.

(4) Group Membership Influences -- Investigations need to be sensitive to the effects of the social structure of the learning setting on individual students. Group composition and students' roles within the groups can affect their opportunities and behavior and the behaviors of significant others (teachers, peers).

(5) Discontinuity Effects -- The continuity of educational experiences of students warrants investigation. Both the students' natural transitions from class to class and the transition from program to non-program participation needs to be traced in order to determine the consequences of discontinuities in instructional and social experiences for educational performance and for the robustness of program influences.

Investigations of the possible impact of school reform which are sensitive to the five features delineated above (interactiveness, process characteristics, implementation, educational and social structure, discontinuity potential) will be comprehensive by necessity. The programs themselves will need to be scrutinized both as proposed and implemented. Intensive and detailed examinations of the characteristics of participants (pupils, teachers, etc.) and educational settings will be required. The experiences of students will have to be followed for a number of years including sufficient time after direct contact with the program has ended.

In essence, we are making a case for viewing appropriate evaluation of school reforms as a systematic body of research. Moreover, given the extensiveness and diversity of the features of program impact, there are good reasons for conducting several overlapping and potentially interlocking studies rather than attempting to embed all features within a single large-scale enterprise. There may simply be too many puzzle pieces for any one group to put together, nor would one group likely have the optimal array of expertise for investigating each individual feature. Finally, there is the potential advantage of

diversity of perspective from mounting multiple smaller investigations as opposed to a single larger one.

The call to treat the evaluation of school reforms as a systematic research agenda involving multiple investigations from diverse perspectives is a recurrent theme in recent writing by some evaluation theorists and methodologists (see, for example, Boruch & Cordray, 1980; Burstein, 1980c; Coleman et al., 1979; Cook & Gruder, 1978; Cronbach, 1978; Cronbach & Associates, 1980; Weiss, 1977). These writers view evaluations as empirical enterprises undertaken in complex political and social settings. Thus, extensive efforts to document the characteristics of the program and multiple methods for examining its consequences are necessary. In their recent book on reforming program evaluation, Cronbach and his associates (1980, p. 72-73) explicitly support this approach to evaluation. They call for a "move away from stand-alone evaluations of programs that address the same social proglems ... also urge that an evaluative effort employ a bundle of studies that use different techniques to examine sub-questions and that the plan be adapted as the studies expose uncertainties more clearly."

Proposed Investigations

While it would presumptuous to propose to delineate all of the multiple investigations for examining the kinds of school reforms considered here, a few of the needed studies can be identified.

Process-Outcome Studies. There should be one or more investigations within a single year of the relationship of entering characteristics to end-of-year performance for students participating in the program. These studies would have to involve several classrooms (preferably at least two classrooms per grade per school from several schools).

Extensive documentation of entering student characteristics, teacher characteristics, the fidelity of the implementation of the program, and the skills and attributes of the student upon leaving the class would be required.

This type of investigation might be modeled after process-product research on teaching (e.g., Far West Lab's Beginning Teacher Evaluation Study; various studies done by the Texas R&D Center; Good and his associates' studies of elementary school mathematics) with the additional wrinkle that specific programs are being investigated and interactions of program properties with student characteristics are anticipated. If the number of programs to be considered gets very large (say, more than 5 or 6), running several such studies each focusing on an overlapping subset of programs (e.g., Study A includes programs 1, 3, 4, 5, and 8; Study B includes programs, 2, 3, 5, 6, and 7; Study C includes programs 1, 2, 4, 6, 7, and 8) would be a reasonable strategy.

These investigations could be used to examine the interactive effects of programs, the effects of program and instructional processes on outcomes, and the influences of the educational and social structures of programs on student performance. The studies are perhaps of less value for examining program implementation, at least as stand-alone evidence. In order to study the discontinuities associated with natural transitions and program-non-program transitions, the studies would have to be extended over several years. The latter would be ideal though there are possibly better investments of resources than maintaining the same level of intensity and size of study over an extended number of years.

Implementation-Process Studies. A second type of investigation would be a possibly overlapping study of the implementation of the various programs. The intent of such studies would be to examine the fidelity of the linkages between program definition and educational processes over a broader set of student, teacher, school, and community characteristics. Besides seeking a broad array of student, teacher, school, and community characteristics, the degree of training provided to implementing sites (schools and classrooms) and the number of years after introduction of the program should be varied. These studies would place less emphasis on measuring student performance and more on measuring the impact of programs on teacher and school practices. They would presumably shed light on the difficulties to be encountered in introducing and disseminating the programs beyond schools participating in the research and evaluation studies.

Longitudinal Studies. Long-term investigations of the patterns of student experiences and performance should also be carried out. These studies would begin with fewer classes of students and follow their specific class and school experiences within programs and for several years after leaving the program. There are some advantages to attaching these long-term studies to the process-outcome studies since extensive documentation would already be available for the first few years.

The long-term study of the pattern of experiences should yield information about the discontinuity effects of various patterns of instructional experiences and program participation. Enough information should be collected about the educational and social setting of class-

rooms containing target students to be able to classify students according to profiles of their educational experiences.

Case Studies of Program Experiences. Some effort should be devoted to documenting the history of children and schools involved in the programs over an extended period of time. A few well-chosen extended case studies of children and schools would provide another form of clarity to the picture of program participation. It might be advisable to begin the case study effort at the same time as the process-outcome investigations so that initial data collection efforts can aid in the selection of prototypic schools and students within schools to follow more intensively.

Exploratory Studies. Finally, a series of "what if" studies should be conducted. One such study could investigate the effects of extending programs up into higher grades rather than expecting them to diffuse naturally or by the influence of students who participated in the program on their post-program teachers. Presumably, the discontinuity effects of program leaving would be weaker under such conditions.

Another "what if" condidate would be an investigation of the changes in the educational processes and performance that might result from removing certain components from the various programs. Such variations as fewer or no aides for programs that are heavily dependent on materials could be tried. In essence, one would be studying the consequences of the program under conditions of diminished resources for its implementation. The data from such a study could contribute to cost-effectiveness studies of specific program features.

Concluding Remarks

Though this paper began with a specific purpose, namely, to consider methodological perspectives on the documentation of school reforms with particular emphasis on the implications of the transitory character of the experiences of program participants, a broad range of topics have actually been discussed at a rather general level. Moreover, there has been no discussion of specific analytical strategies for documenting program impact.

The explanation for the apparent shift from the intended purpose can be traced to the nature of the school reform effort and what is known about how to document it. On the one hand, it can be argued that past research and evaluations provide a lot of information about specific aspects of school reforms when the confounding of elements of the reform and the limitations of conceptualization, design, instrumentation, and analysis are ignored. On the other hand, it is clear that elements of school reform are confounded and the various limitations should not be ignored. Program elements are inherently interrelated and their interface, linkages, and dependencies are at the heart of a sound understanding of school reform efforts. Better conceptualization, design, instrumentation, and analyses are possible in further investigaions. But the resulting improvements will be marginal at best unless the refinements are directed toward understanding both program elements and their interrelationships.

To this end, we argue that the transitory character of the experiences of program participants is simply an inherent feature of the educational processes and outcomes of educational innovations. Thus, we argue

that the impact of student transitions and varying group membership should be examined within a conceptual framework that allows consideration of a full range of features. Each feature -- identified here as interactiveness, process characteristics, implementation, educational and social structure, discontinuity potential -- reflects to a greater or lesser degree the effects of student transitions and varying group memberships.

If there is to be another round of research and development on reforms in early education, then its thrust and perspective should be broadened to encompass a fuller range of experiences and outcomes than in past investigations. This effort should also be cognizant of the essential dynamic properties of educational change and consequently, the central role of educational and social processes in school reform. On these grounds, a research strategy focusing on educational processes and experiences -- as intended and observed outcomes of school reform programs and as antecedents of educational performance -- has been recommended. Combining the focus on educational and social processes with multiple investigations from diverse perspectives offers the opportunity for valuable information for informed decision-making and a more realistic view of school reforms and their accomplishments.

References

Anderson, R.B., St. Pierre, R.G., Proper, E.C., & Stebbins, L.B. Pardon us, but what was the question again? A Response to the Critique of the Follow Through Evaluation, Harvard Educational Review, 1978, 48(2), 161-170.

Armbruster, B.B., Stevens, R.O., & Rosenshine, B. Analyzing content coverage and emphasis: A study of three curricula and two sets. Technical Report No. 26, Center for the Study of Reading, University of Illinois, Urbana-Champaign, 1977.

Baker, E.L. Evaluation of the California early childhood education program, Vol. 1, Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles, 1976.

Bidwell, C.E. & Kasarda, J.D. Conceptualizing and measuring the effects of school and schooling. American Journal of Education, 1980a, 88, 401-430.

Bidwell, C.E. & Kasarda, J.D. Problems of multilevel measurement: The case of school and schooling, In K.H. Roberts & L. Burstein (Eds.) Issues in Aggregation, No. 6, New Directions for Methodology of Social and Behavioral Science, San Francisco, Calif.: Jossey-Bass, Inc., 1980b, 53-64.

Boruch, R.F. & Cordray, D.S. An appraisal of educational program evaluations: Federal, state and local agencies, Washington, D.C.: U.S. Department of Education, 1980.

Bossert, S.T. Tasks and social relationships in classrooms. New York: Cambridge University Press, 1979.

Brown, B. Long-term gains from early intervention: An overview of current research. Paper presented at the meeting of the American Association for the Advancement of Science, Denver, February 1977.

Brown, B.W. & Saks, D.H. The production and distribution of cognitive skills within schools. Journal of Political Economy, 1975, 83, 571-593.

Brown, B.W. & Saks, D.H. Production technologies and resource allocations within classrooms and schools: Theory and measurement. In R. Dreeben & J.A. Thomas (Eds.) The analysis of educational productivity, Vol. I: Issues in microanalysis. Cambridge, Mass.: Ballinger, 1980a.

Brown, B.W. & Saks, D.H. Economic grouping and data disaggregation. In K.H. Roberts & L. Burstein (Eds.) Issues in aggregation, No. 6, New directions for methodology of social and behavioral science, San Francisco, Calif.: Jossey-Bass, Inc., 1980b, 31-42.

Burstein, L. The role of levels of analysis in the specification of educational effects. In R. Dreeben & J.A. Thomas (Eds.), The analysis of educational productivity, Vol. I: Issues in microanalysis. Cambridge, Mass.: Ballinger, 1980a, 119-190.

Burstein, L. The analysis of multilevel data in educational research and evaluation. In D.E. Berliner (Ed.) Review of research in education, Vol. 8, Washington, D.C.: American Educational Research Association, 1980b, 158-233.

Burstein, L. Paired/parallel evaluations: The potential of multiple investigations in large-scale program evaluation. Presented at the International Consortium on Education Evaluation (May 12th Evaluation Group) Meeting, Greensboro, North Carolina, November 1980c.

Calfee, R. & Brown, R. Grouping students for instruction. In D.L. Duke (Ed.) Classroom management. 78th Yearbook of the National Society for the Study of Education, Part II, Chicago: The University of Chicago Press, 1979, 144-181.

Cicirelli, V.G. et al. The impact of Head Start: An evaluation of the effects of Head Start on children's cognitive and affective development, Vol. 1 & 2. (Report to the Office of Economic Opportunity, Contract B89-4536). Ohio University, Athens: Westinghouse Learning Corp., New York, N.Y., June 1969.

Cline, M.G., Ames, N., Anderson, R., Bale, R., Ferb, T., Joshi, M., Kane, M., Larson, J., Part, D., Proper, E., Stebbins, L., & Stern, C. Education as experimentation: Evaluation of the Follow Through Planned Variation odel, Vol. 1a, 1b. Cambridge, Mass.: Abt Associates, 1974.

Cohen, D.K. The value of social experiments. In A.M. Rivlin & P.M. Timpane (Eds.) Planned variation in education: Should we give up or try harder? Washington, D.C.: The Brookings Institute, 1965, 147-168.

Coleman, J.S., et al. Policy issues and research design. Chicago: National Opinion Research Center, 1979.

Cook, T.D. & Gruder, C.L. Metaevaluation research, Evaluation Quarterly, 1978, 2(1), 5-51.

Cronbach, L.J. Designing educational evaluations, Occasional Paper, Stanford Evaluation Consortium, Stanford University, 1978.

Cronbach, L.J., Ambron, S.R., Dornbusch, S.M., Hess, R.D., Hornik, R.C., Phillips, D.C., Walker, D.F., & Weiner, S.S. Toward reform of program evaluation, San Francisco, Calif.: Jossey-Bass Publishers, 1980.

Cronbach, L.J., & Snow, R.E. Aptitudes and instructional methods, New York: Irvington Press, 1977.

Datta, L-e. Design of the Head Start Planned Variation experiment. In A.M. Rivlin & P.M. Timpane (Eds.) Planned variation in education: Should we give up or try harder?, Washington, D.C.: The Brookings Institute, 1975, 79-100.

Ellett, C.D., Haun, H.C., Pool, K.W., & Smock, C.D. Planning information study for future Follow Through experiments: Task I response general issues, Mathemagenics Activities Program, University of Georgia, Athens, Georgia, May 1979.

Firebaugh, G. Groups as contexts and frog ponds. In K.H. Roberts & L. Burstein (Eds.) Issues in aggregation, No. 6, New directions for methodology of social and behavioral sciences, San Francisco: Jossey-Bass Inc., 1980, 43-52.

Haney, W. A technical history of the National Follow Through evaluation. Cambridge, Mass.: Huron Institute, 1977.

Haney, W. Units and levels of analysis in large-scale evaluation. In K.H. Roberts & L. Burstein (Eds.) Issues in aggregation, No. 6, New directions for methodology of social and behavioral sciences, San Francisco: Jossey-Bass, Inc., 1980, 1-16.

House, E.R., Glass, G.V., McLean, L.D., & Walker, D.F. No simple answer: Critique of the Follow Through evaluation. Harvard Educational Review, 1978, 48(2), 128-160.

Johnson, D.W. Student interaction: The neglected variable in education. Educational Researcher, 1981, 10(1), 5-10.

Kennedy, M.M. Findings from the Follow Through Planned Variation study. Educational Researcher, 1978, 7, 3-11.

Lazar, I., Hubbell, V.R., Murray, H., Rosche, M. & Royce, J. The persistence of preschool effects: A long-term follow-up of fourteen infant and preschool experiments. Final Report, Consortium on Developmental Continuity, Ithaca, New York, September 1977.

Madaus, G.F., Airasian, P.W., & Kellaghan, T. School effectiveness. New York: McGraw-Hill, 1980.

Lukas, C.V. Problems in implementing Head Start Planned Variation models. In A.M. Rivlin & P.M. Timpane (Eds.) Planned variation in Education: should we give up or try harder? Washington, D.C.: The Brookings Institute, 1975, 113-126.

Palmer, F.H. The effects of early childhood educational intervention on school performance. Paper prepared for the President's Commission on Mental Health, Washington, D.C., July 1977.

Porter, A.C., Schmidt, W.H., Floden, R.E., & Freeman, D.J. Practical significance in program. American Educational Research Journal, 1978, 5(4), 351-357.

Rivlin, A.M., & Timpane, P.M. Planned variation in education: An assessment. In A.M. Rivlin & P.M. Timpane (Eds.) Planned variation in education: Should we give up or try harder? Washington, D.C.: The Brookings Institute, 1975, 1-22.

Rivlin, A.M., & Timpane, P.M. (Eds). Planned variation in education: Should we give up or try harder? Washington, D.C.: The Brookings Institute, 1975.

Scan Sponsor Task Force. Follow Through: Forces for change in the primary school. Ypsilanti, Michigan: The High/Scope Press, 1980.

Sharan, S., Ackerman, Z., & Hertz, Lazarowitz, R. Academic achievement of elementary school children in small-group versus whole-class instruction. The Journal of Experimental Education, 1978/80, 48(2), 125-129.

Smith, M.S. Some short term effects of Project Head Start: a report on the second year of planned variation--1970-1971. Cambridge, Mass.: The Huron Institute, 1973.

Smith, M.S. Evaluation findings in Head Start Planned Variation. In A.M. Rivlin & P.M. Timpane (Eds.) Planned variation in education: Should we give up or try harder? Washington, D.C.: The Brookings Institute, 1975, 101-111.

Stallings, J. Implementation and child effects of teaching practices in Follow Through classrooms. Monographs of the Society for Research in Child Development, Serial No. 163, 1975, 40(7-8).

Stebbins, L.B., St. Pierre, R.G., Proper, E.C., Anderson, R.B., & Cerva, T.R. Education as experimentation: A planned variation model, Vol. IV-A, An evaluation of Follow Through. Cambridge, Mass.: Abt Associates, 1977.

Walker, D.F., & Schaffarzick, J. Comparing curricula. Review of Educational Research, 1974, 44, 83-111.

Wargo, M.J., & Green, D.R. (Eds.) Achievement testing of disadvantaged and minority students for educational program evaluation. Monterey, Calif.: CTB/McGraw-Hill, 1978.

Webb, N.M. Group process: The key to learning in groups. In K.H. Roberts & L. Burstein (Eds.) Issues in aggregation, No. 6, New directions for methodology of social and behavioral sciences. San Francisco: Jossey-Bass Inc., 1980, 77-88.

Weikart, D.P., & Banet, B.A. Model design problems in Follow Through. In A.M. Rivlin & P.M. Timpane (Eds.) Planned variation in education: should we give up or try harder? Washington, D.C.: The Brookings Institute, 1975, 61-78.

Weiss, C.H. (Ed.) Using social research in public policy making. Lexington, Mass.: Lexington Books, 1977.